



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2002

Personalizing Textbooks with Slicing Technologies - Concept, Tools, Architecture, Collaborative Use

Schwabe, Gerhard ; Dahn, Ingo

DOI: <https://doi.org/10.1109/HICSS.2002.994542>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-67788>

Conference or Workshop Item

Originally published at:

Schwabe, Gerhard; Dahn, Ingo (2002). Personalizing Textbooks with Slicing Technologies - Concept, Tools, Architecture, Collaborative Use. In: Hawaii International Conference on System Sciences (HICSS02), Hawaii, 1 January 2002.

DOI: <https://doi.org/10.1109/HICSS.2002.994542>

Personalizing Textbooks with Slicing Technologies – Concept, Tools, Architecture, Collaborative Use

Ingo Dahn, Gerhard Schwabe, University of Koblenz-Landau, Germany
dahn|schwabe@uni-koblenz.de

Abstract

Electronic books are not yet heavily used because they still do not offer sufficient benefits. The Slicing Book approach is a new approach to electronic books offering personalized electronic textbooks for learners. Books are split into slices and the logical relationship between those slices is made explicit, e.g. what slice builds on knowledge provided by what other slices. This knowledge can be matched with a model of the user's knowledge to provide him or her a personalized textbook. The paper describes the basic concepts of slicing technologies, tools and architecture and the authoring process and usage scenarios. Slicing technologies have the potential to change books' roles in teaching. This is demonstrated in the domain of collaborative learning.

1 Introduction

Currently – in the year 2001 – the market for electronic books has not taken off yet. The alternative to read on a screen what is also available in print did not convince a supercritical mass of customers to buy dedicated reading devices.

Competing with printed books, electronic books are evaluated by the potential customers according to four criteria.

1. Availability of features that they know from printed books, like ease of reading, typographic quality, annotations, independence of additional devices etc.
2. Availability of additional features that they do not know from printed books, like advanced search options, multimedia content, links to other materials, personalization etc.
3. Availability of interesting content.
4. Availability on the market at a competitive price.

Finally the balance of these four criteria will decide on the market success of electronic books.

E-book readers which are nowadays on the market still have technical limitations that prevent them from offering the comfort that printed books offer. Here it may suffice to mention the lack of advanced typographic features and of colour. However, electronic documents currently

available on laptop and desktop computers show, that these restrictions can be overcome. There is no fundamental reason to be seen why the power of a modern desktop computer cannot be compressed into a device of the size and weight of an average book.

The argument that the feel and smell of a newly printed book cannot be imitated by an electronic device is valid in the short run since average people prefer to buy products they are already used to. However, this deficiency can be outweighed in the long run by the availability of additional features, leading to a new culture of information exchange.

In fact, such cultural changes have occurred in the past several times. The preferred form for the exchange of knowledge and novels has evolved from oral communication through handwritten manuscripts to the current mass production of books. Gutenberg's books were not as beautiful as the books handwritten by medieval monks, but they found their market because they offered the additional feature of being widely available at an affordable price.

Similarly in the modern world the availability of new attractive features at an affordable price will determine the future role of electronic books. In this paper we shall not discuss the various forms that a transition from printed books to electronic books – possibly in combination with CD-ROM – may take. Some considerations on these aspects will appear in [Da00a].

Instead we shall describe a technology that equips electronic documents with an important feature not found in printed books – personalization. Moreover – since we are aware that cost efficiency is critical for market acceptance – we shall describe how already existing documents designed for print production can be re-engineered to offer personalization services in an electronic form.

Our report is based on experience gained within the project TRIAL-SOLUTION (<http://www.trial-solution.de>) supported by the European Commission, the project IN2MATH (<http://www.uni-koblenz.de/ag-ki/>) supported by the German Ministry for Education and Research and SYWIKOL supported by the Deutsche Forschungsgemeinschaft. The first of these projects emphasizes the generation of personalized documents from various sources, the second adds interactive

components to these personalized documents and investigates the authoring of documents suited for personalization and the third is concerned with aspects of collaboration in the authoring and use of personalized documents. The approach described has led to the publication of two electronic books by Springer (publications by other publishers are in preparation) and to their adoption for teaching purposes in several German Universities. Furthermore we are in the process of an explorative evaluation of advanced features in the lecture on logics.

The considerations below apply to all kinds of well-structured documents. We shall however concentrate on the generation and use of personalized electronic documents for educational purposes. We regard the learning domain as particularly interesting because

a) here books are widely available and heavily used; the domain is therefore an interesting market for publishers,

b) books are embedded in an educational process; improving books' characteristics has the potential to improve their role in the learning process and ultimately may lead to improved or more efficient learning. The application in the learning domain may therefore lead to interesting effects. We will explore this potential in the domain of collaborative learning.

Section 2 will introduce the basics of the slicing book technology, Section 3 will present the tools necessary to develop and use slicing books and the basic system architecture. Sections 4-6 describe the application side of slicing books: Section 4 focusses on the side of the author, section 5 synthesizes experiences and ideas into usage scenarios and section 6 introduces the idea of collaborative slicing. The final section 7 compares the slicing book approach to related concepts.

2 Slicing Book Technology

Already hypertexts offer readers new qualities of service. Links take readers instantly from any point to any other point which is referenced at this spot. Thus the classical hierarchical document structure is augmented or even replaced by a network of pages. However still information that is needed for a specific purpose remains scattered within this network. Slicing Book Technology goes one step further by contracting this information into a new document. Thus it can be seen as a combination of hypertexts with the dynamic generation of documents.

A sliced book consists of

- a hierarchy of semantic units together with
- a set of meta-data.

Unlike web pages, semantic units need not be complete documents that are ready for delivery. Rather they can be pieces of information that have to be combined with other pieces and processed appropriately in order to obtain a

deliverable document. For example to obtain a deliverable HTML page a number of fragments of HTML code may be combined into a new file, augmented with HTML header and footer and processed with an appropriate style sheet. This process that leads to the generation of a new page may be varied in many ways. Especially it may depend on characteristics of the particular user who has requested that page or whether a certain fragment is included or not. Also the particular style sheet may be chosen to suit the users preferences.

This offers many possibilities for the production of personalized documents. In order to select among these possibilities, meta-data are used. They contain for each of the semantic units the information about this unit and about its relation with other units that an intelligent system may need in order to decide whether this unit should be included in a requested document. For example these meta-data for a unit E may state that E is an exercise on topic T that requires the prerequisite units P and Q for its solution. Then a teacher looking for exercises for an exam on topic T may get delivered just E but learners may get delivered also P and Q in order to prepare for that exam.

Note that the knowledge how to compose documents for teachers or learners can be formulated in a general way, i.e. independent of the concrete units. Appropriate knowledge management procedures combine this general knowledge, the meta-data describing the content and the available information about the users and their current interest. They infer, which slices should be presented to the user and how the document should be formatted. If possible, the knowledge management will resolve conflicts and select among possible alternatives.

Finally, based on the recommendations of the knowledge management system, a new document will be generated on the fly and delivered to the user. We note that the described architecture consisting of a collection of meta-data annotated slices, a knowledge management system and a dynamic generation of personalized documents, coincides with the vision of the future World Wide Web as a semantic net. This vision was stated by Tim Berners-Lee – the inventor of the World Wide Web – at the conference XML 2000.

Since May 2000 the first server using Slicing Book Technology is online at <http://www.slicing-infotech.de/slib/analysis01-en.html> [WD00]. A server with content in English goes online at <http://www.slicing-infotech.de/slib/logic01.html> [Fu01] in June 2001.

In order to make Slicing Book Technology work, collections of interoperable annotated slices must be available. Such collections can be either created from scratch or they can be generated by the adaptation of existing documents. Subsequently we shall concentrate on these authoring processes. Our report is based on the experience we gained during the authoring of [WD00],

[Fu01] and during the preparation of further pre-existing books within the Trial-Solution project.

3 Slicing Books – Tools and Architecture

The production of sliced books, either from scratch or from existing documents, needs to take different processes into account.

1. The production of the content (data) in an appropriate format
2. The assignment of meta data describing these data
3. The description of the desired layout respectively the desired layouts
4. The description of the users, their interests and preferences
5. The acquisition of the general knowledge which documents are to be generated for users with specific user models

None of these can be totally automated, but – since the results of these processes are to be processed automatically – machine support would be desirable for all. Fortunately, not all of these processes have to be repeated for each sliced book. Especially the components of a user model and the description of the general directives for the generation of documents (i.e. items 4 and 5) can be designed once for a number of books from the same domain.

Textbooks are not interactive. Therefore, sliced books that are generated from existing textbooks will have only very limited information on the interests and knowledge of the user. Unlike computer based training systems they will lack an assessment component that can provide reliable information on the knowledge that the user has acquired. Even if textbooks contain exercises, it is frequently impossible to verify automatically the correctness of the solution that a student might provide. Therefore, the user model of a sliced book can be kept as simple as an overlay model (see [CG77]) which only records what the users know and what they want to know.

Using techniques from artificial intelligence, notably from automated deduction, the description of the documents to be generated can be given in a descriptive rather than an algorithmic form. Such descriptions consist of rules like the following.

If the user is a learner and looks for an exercise on topic T then from each group of such exercises two exercises are included into the document. Moreover all prerequisite units for these exercises are included.

In addition there may be other rules required in order to complete the description, for example:

If n units are to be selected and there are more than n units available for selection, then select those units which are closer to the beginning of the book but are not yet known to the reader.

In our experiments we found that quite a small number of rules (3 – 20) – in combination with thousands of simple facts describing the assignment of meta data to the units – are sufficient to provide useful personalized documents. This reduces the need to provide special editors for the description of documents.

Tools for handling item 3 of the above list – the description of the presentation style – are readily available on the market in the form of standardized styles for various types of documents, especially for XML, HTML and LaTeX documents. When editing XML documents becomes better supported by popular word processors standard XML styles will play for all scientific documents the same role that LaTeX styles have taken over the last decade in the field of mathematics and natural sciences. These styles are in most cases provided by publishers. They have proved to be an efficient way to obtain a well designed and unified layout that can be easily adapted to different needs.

So we are left with the need to provide tools to support the generation of the content and the assignment of meta data in order to get the production of sliced books going. For the reengineering of existing documents such tools are currently under development in the Trial-Solution tool set. Before we can describe this tool set, we have to mention another important feature that makes personalized electronic books potentially superior to printed books.

Currently information is preferably exchanged in the form of static documents, in electronic or printed form. This explains why information retrieval and document retrieval are frequently regarded to be identical. However, this is no longer adequate. Already XML splits the document into content and presentation (style). Slicing Book Technology goes one step further and frees the information from its fixed position in the content of a particular document. It supports the aggregation of content for new documents using meta data (Figure 1).

Given this situation, electronic books can be generated from parts of different documents, for example in order to make use of the specific expertise of different authors on particular aspects of a problem. In order to support this, it is necessary that the meta data of all units fit into a coherent meta data system. On the other hand, the meta data system needs to be fine enough to differentiate between the semantic units to be combined. Semantic units differ mainly by their content. Hence a refined content description system is required and this system has to be kept coherent. The Trial-Solution tool set takes these problems into account by using automated key phrase assignment and a central key phrase server.

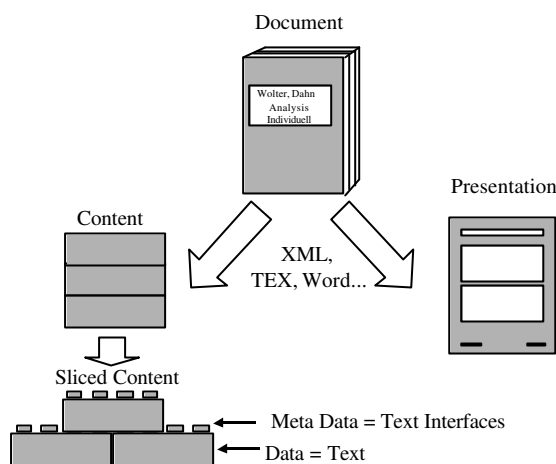


Figure 1: Separation of Documents in Presentation and Content and Subsequently Sliced Content

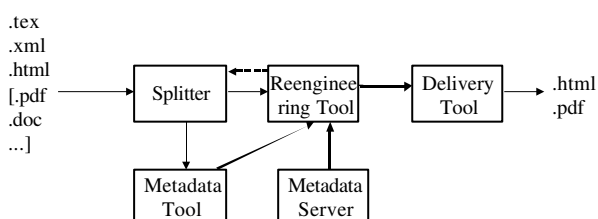


Figure 2: System Modules

When an existing document is to be sliced, it is first disaggregated automatically by the Splitter (Figure 2). The Splitter also extracts meta information that is explicit in the document, like references, key phrases assigned by the author or semantic types of units. It takes structural information, formatting and specific word usage into account and can be adjusted to the writing style of a specific author. The Splitter has been developed within the TRIAL-SOLUTION project by Slicing Information Technology Berlin (SIT). The Splitter private edition that handles LaTeX documents can be downloaded from the project's web site.

In realistic documents not all required information can be extracted automatically. Therefore, in the next step, the sliced document is fed into the automatic key phrase assigner. This system assigns key phrases from a controlled vocabulary to parts of this document using the method of enriched thesauri [Ha97]. Note that semantic units can be quite small, sometimes they consist of a single line. Therefore they may not contain enough

linguistic material for a reliable key phrase assignment. In such a situation key phrases will be assigned to groups of slices and inherited by the single slices from these groups. Moreover the key phrase assigner classifies (groups of) units according to a domain specific classification system.

The resulting enriched document is fed into the Reengineering Tool. This tool contacts a central key phrase server to ask which key phrases have been assigned already to slices from other books with the same classification. Thus the user of the Reengineering Tool is provided with a collection of key phrases that are of potential relevance for the sliced book to be reengineered.

Using the reengineering tool the following meta data are revised manually.

- Assigned key phrases. The Reengineering Tool is also used to define new key phrases when no appropriate key phrases have been proposed.
- Types of slices (like “example”, “algorithm” etc.)
- Relations between slices (like “which slices are pre-requisite to the given slice”)
- Titles that can be used to navigate between slices

Also the structure of the document as recognized by the Splitter may require a manual revision with the Reengineering Tool. Sometimes, especially when the author has changed its writing style within the book, it will be necessary to re-slice parts of the document with different settings. To enable this, a partially reengineered document can be sent back from the Reengineering Tool to the Splitter for revision.

All this requires domain specific competence and is best done by the authors themselves. However it can be also completed by students who have mastered the subject of the book.

Domain experts are not experts in thesaurus design. The key phrases they have newly defined are sent to the central Key Phrase Server where they are integrated by competent personnel into the overall thesaurus. Then they are either confirmed or a file will be sent back to the Reengineering Tool that controls the automated replacement of assigned key phrases by others that have the same meaning.

Finally, when the reengineering is completed, the sliced book is sent to the Delivery Tool. This consists of a server that takes user requests and builds personalized documents from the slices of the books that it has imported.

All these tools exchange their data in the Content Packaging Format specified by the IMS Global Learning Consortium (<http://www.imsproject.org>). Meta data and thesauri are also exchanged in XML-defined formats.

A description of these formats would go beyond the scope of this article. Also other relevant aspects (e.g. potential business models, copyright issues or the problems stemming from the generation of new books

from books that have already been personalized) cannot be discussed here. Instead we come now to discuss the “human side” of the technology.

4 The Authoring Process

Designing materials specifically for electronic media is currently rather expensive. It requires a combination of competences that hardly any author can provide. On the other hand there is a large number of approved and appreciated documents available in print. In this section we will describe the process that makes these documents fit for online delivery, personalisation and added value electronic services from an author’s point of view.

For use with Slicing Book Technology an existing document must be disaggregated into a set of slices and these slices must be augmented with meta-data. Since slices are intended for reuse, the content should be organised in a well-structured way so that the isolation of reusable units is possible.

Good textbooks, legal and technical documents are examples for which this requirement is frequently met. It is important that start and end of each unit are clearly marked. It is not important how this is done if only it is done throughout the document in a uniform way. For each type of unit start and end can be marked in specific ways. This can even help to detect the type of a unit automatically. Formatting is also a possible way to determine units. For example, when definitions start with the word “Definition”, are printed in italics and the defined concept is printed in bold, then definitions can be automatically isolated and the defined concept can be extracted as a candidate key phrase. On the other hand, if the definition is not specially formatted, consists of several paragraphs and is followed by an explanatory text in another paragraph it may be impossible to detect the end of the definition automatically.

In such cases the missing end can be manually augmented in two ways. One way is to insert additional marks that can be detected during the splitting process. Ideally this is done by invisible marks which do not appear in print or screen representation. The other method is to use the Reengineering Tool to restructure the document and to split a part of text into units of the right size.

The structure of the book should clearly reflect its logical structure. A bad example, found in one of the books to be sliced, is a group of 5 exercises preceded by an introductory sentence “*In the following equations determine x and y* ”, where y occurred only in the first two equations whereas the other 3 equations contained x only.

The personalized documents that are generated from slices will be composed from slices of different size. Also the formatting may be adapted to the users needs. Therefore, documents in a data format that links their contents intimately with their layout are hardly

appropriate. We mention QuarkXPress and Postscript here. XML-based documents are ideal candidates since they are clearly structured according to a Document Type Definition and have the layout separated in their corresponding style sheets. LaTeX documents are appropriate and Microsoft Word documents can be used.

When the author has provided a hand-made design that optimises the document for print on a certain paper size, all these optimisations have to be replaced by more general solutions that are also valid for screen rendering in the various combinations that might occur. To achieve this automatically, it is very helpful to use styles (which the author may have even defined himself) instead of formatting parts by hand.

A bad example, found frequently in MS Word manuscripts is, that paragraphs are separated by blank lines, i.e. there occur additional paragraphs that do not have content. This demonstrates that—until now—many authors do not distinguish clearly between layout and content. The same effect could be easily achieved by defining the style of the paragraph in an appropriate way. Then, by replacing this style definition, the text could be easily reformatted to make better use of the screen size or to adapt to the style of units from another document.

As a rule of thumb one can state that the less effort the author has invested into the layout, the better are the chances that the slices can be reformatted for reuse as needed.

There are also cases where automated reformatting is impossible. For example figures are very interesting objects for reuse since their production can be expensive. In textbooks, figures sometimes come with a lot of text that explains them. Now, when the user wants to reuse the figure formatted as an overhead slide, there is no way to let the explanatory text fit on the slide together with the figure. In this case the author has to foresee this intended reuse and isolate the figure (eventually with a smaller explanatory text) from the bulk of the text.

Before a document will be sliced, the required granularity has to be decided. The size of a slice may vary between a chapter and a single line. The more fine grained the structure is, the more flexible possibilities for the generation of documents are obtained. However a more fine-grained structure also results in a larger set of slices and requires more meta-data to be added for the description of these slices. Therefore the intended quality of service and also economic considerations determine which parts of a document are stored as separate slices. The following definition may help to determine appropriate parts (see [Da01]).

Definition. A potential slice is a connected part of a document that can be reused under well defined conditions.

Here “well defined” means that conditions for reuse can be stated explicitly in a formal way so that they can be processed by an automated advisory system.

It has turned out that unadvised automated slicing may deliver too fine-grained results. Therefore as the first step of the reengineering process it should be determined by the author or by another domain expert what are the units that are interesting for reuse and how they can be determined so that the Splitter can be configured in an optimal way. The fine-grained mathematical textbooks handled in the TRIAL-SOLUTION project have a between 1.300 and 3.000 slices for each book of about 450 pages. The size of the slices varies between a single line (for example for some exercises), and a few pages (for example proofs).

For the manual reengineering it is important to understand that sliced books can be conceptually interpreted as a tree of files. Leafs in this tree are atomic slices while inner nodes represent aggregations of slices. Each node may have a file with its specific content and two more files, named start- and end-file. The purpose of these start- and end-files is to ensure that generated documents are technically well formed.

To achieve this, whenever the content of a node is to be included into a personalized document, this content will be enclosed between the start- and end-files of all parent nodes of the given node. For example, if a list in an HTML document is to be disaggregated, the list may be represented by a node that has a descendant for each item of the list and the start- and end-tags of the list in its start-respectively end-file. In this way, whenever a list item is to be presented, it will be enclosed between the appropriate list tags.

When this tree of files is traversed in a depth-first way, the original document is reconstructed. This is an important feature since it makes sure that the original work of the author is only augmented and reused but not changed.

The Reengineering Tool supports the restructuring of the document but it does not permit changes in the sequence of the content. Thus content can be only moved into neighboring slices. New groups of slices can be created or resolved and slices can be split or joined. These operations may corrupt locally the soundness of the structure and their results must be checked.

Adding relations between slices is one of the most important functions of the reengineering process. In this way, prerequisite knowledge necessary to understand a given slice is recorded. This can be done in two ways.

1. Concept oriented: It is recorded which concept (represented by a key phrase) is prerequisite to which other concept (key phrase)
2. Content oriented: It is recorded which units are assumed to be known in order to understand a given unit.

We have chosen the second approach since it is more refined. Relations between key phrases can be derived

from the recorded relations between units that have these key phrases assigned. This approach also has turned out to be more comfortable for the author who is used to work directly with the content.

Assignment of key phrases and recording relations between units is done in parallel. Thus the tool can propose key phrases that have been assigned to related units or it can propose potentially related units that have similar key phrases assigned.

The meta-data described so far are the basis to determine which slices are to be presented to a reader in a concrete situation. Other meta-data can give the author and the publisher the opportunity to control how this is done.

Most important is the inclusion of information that describes intellectual property rights. Especially it must be possible to trace the authors, source and original location of each slice included into a personalized document. This kind of meta-data is usually inherited from the description of the complete document. The author or publisher may also wish to restrict the re-use of a set of slices. For example they may not allow the rearrangement of slices or the combination with material of a competitor. This information is not contained in printed books and hence cannot be extracted automatically.

The meta-data system described here can be extended in various ways if this is useful to provide additional services. Some international consortia, for example IEEE, Dublin Core and IMS, have specified their own meta-data systems. None of these extensions is well-suited for automated meta-data assignment. Especially educational meta data that describe for example the target audience or how long learning a specific unit will take are hard to define in a sufficiently standardized way. For this reason they are not used in the TRIAL-SOLUTION project. Instead, the readers will be given the possibility to express preferences between documents or to select specific types of slices for the generation of their personalized documents.

The Trial-Solution project decided to develop a new meta data system specification since a need was seen for meta data inheritance along the document structure and for refined content descriptions using key phrases from a structured thesaurus (ontology). For meta data related to intellectual property rights, the appropriate part of the IMS Meta Data Specification has been adopted.

A special problem in the composition of personalized books from several sliced sources is to ensure, that all relevant meta data carry over from the source document into the personalized document so that the source of each piece of content can be tracked and that also the creator of the new composition is recognized. These goals are achieved by using an export mechanism that exports together with each slice always the complete path leading

from the document root to this slice with all meta data attached to this node. Then these paths are connecting by a new top node. The meta data attached to this top node keep the information about the creator of the composed document. Note that this procedure can even be iterated so that personalized books can be used as new sliced books from which more (secondary) personalized documents can be created.

Authoring becomes more complex if the document is not preexisting but created and used as a sliced document in parallel or if a sliced document is to be modified. The Reengineering Tool offers only rudimentary editing facilities. On the other hand, editors do not respect meta data that have already been assigned and do not support fine grained structuring of the document. Thus the author is forced to use the editor in parallel with the Reengineering Tool.

A collaborative approach where one author writes the content and another assigns the meta data can simplify this. We are currently also experimenting with an integration of the Reengineering Tool and the EMACS editor.

5 Usage scenarios and experiences

Slicing Book Technology has apparently an unlimited number of usage scenarios. It provides only raw material and the concrete applications determine the ways in which personalized documents are composed. Here we shall mention only a few possibilities that make sliced books superior to printed books and that are possible with the sliced books currently available. The usage scenarios are based on the usage of slicing books in the context of the Trial-Solution project.

5.1 Searching

Besides full text search, which is nowadays standard in electronic publications, sliced books offer advanced key search possibilities. The reader can search for a concept and can take related concepts or sub-concepts into account.

Supplying the part of a key phrase only or using a synonym instead of the exact key phrase will not be an obstacle for a successful search. In these cases where the reader does not know all the available key phrases, the system will show him what might apply. Search can be restricted to deliver only units of specific types. For instance the reader may search for examples only.

This is an efficient way to locate needed information. Frequently the search results are then further adapted by restricting the search or by automatically adding the necessary prerequisites.

5.2 Reformatting

The layout of the document may be adapted to specific needs. Especially the formatting as overhead slides has turned out to be very useful. The teacher locates the parts that he wants to present and collects them in his personal book on the server. Slices that contain only connecting text are removed. Then a document is generated where the remaining slices are automatically formatted as overhead slides.

These slides are either directly projected from the teacher's computer or they are printed.

5.3 Information sharing

The reader may request the selection for her personal book and what she has declared to be known content from the server. She will receive this information in the form of an HTML page that she can store on her own computer. This page does not contain the content itself, but only a table of content.

Moreover this page contains a form through which one can log into the server and add the content referred on the page to one's own personal book. Thus a teacher may provide the recommended reading for the next seminar as such a page on her web site.

Students may log into the server from the teacher's page and her recommendations will be automatically copied into their personal book. Then they may adapt the selection further (for example by adding more prerequisites) and generate their personal preparatory manuscript for the seminar. Especially they might have the slides that the teacher uses but formatted for print or for their laptop screen.

It turns out that users have difficulties to realize the possibilities that the exchange of tables of content (meta data) instead of the content (data) itself offers. For example, some students prefer to receive a file that contains exactly the slides that the teacher has used. They need help to realize that they may determine the formatting themselves.

5.4 Course planning

Rarely, a course reuses a book completely. Usually the teacher will select some specific units that she has to teach. Then she looks up the necessary prerequisites in the literature. These prerequisites require other prerequisites etc. At some point the teacher will reach prerequisites that he assumes to be known. From these on she will work forward in her course.

This process can be automated when sliced books are used, based on the acquired meta data. In practice the teacher will select a few learning objectives and will let the server do the iterated addition of prerequisites. The

server will also add connecting text etc. according to its rules of document composition. Frequently the resulting document will be too large since it contains material that is supposed to be known.

Then the teacher may mark a few of the selected slices as known and the server will infer according to its rules what else must be known if these particular slices are known. Then it will remove all slices from the document that are inferred as known. This process can be iterated and the teacher may also manually add or delete slices. The final result is a course specific manuscript that leads from the known slices to the learning objectives without detours.

The same procedure, called script generation, is also useful for students that want to prepare for a specific question in an exam and need some recommended reading. Script generation is an innovative feature that is totally unexpected by the reader.

Readers expect that they have to select from a book what they want to read, not that the book itself proposes something and takes the reader's knowledge into account. They are used to references in books that take them to closely related parts, but they are not prepared to be taken to distant parts from where on they should start studying in order to understand what they want to learn. This reflects that many students tend to concentrate on details and that they have difficulties to understand the overall picture of a domain.

5.5 Forming a team

In this scenario a group of people comes new into a project. Each participant has his or her particular expertise but has to acquire specific additional knowledge in order to play his or her part in the project and to communicate with the other members of the team.

The team leader can prepare for each new member a particular script for reading – as described in the last part – that avoids what the person knows already and leads to what he specifically needs to know. It may be necessary to collect slices from various books to present all required information and these books may have overlapping content. In this scenario the rules for the generation of documents should be set up such that the scripts for all members of the team use the same slices when they have to provide the same information in order to facilitate communication between the team members during the work.

6 Collaborative Slicing

The usage scenarios presented up to now are based on the notion that a slicing book is an augmented book, i.e. a text is prepared by a set of (knowing) authors and consumed by a group to (less knowing) students. The

advantage in comparison to traditional books is the personalization of the presented content. However, this approach has some limitations:

1. The presented material has to contain a superset of knowledge required for understanding the topic. Otherwise, if frequent queries for explanations are not successful, the learner becomes frustrated. Slicing books make it easier to combine material from different sources, but the necessity of completeness is still there.

2. Authors are detached from the learners. Traditionally, textbooks are used by a disjunct triad of authors, teachers and learners. Slicing books allow the teachers to adapt books to their courses and combine them for their courses. But preparing additional content would require the knowledge and – more important – the time and precision of a (slicing) book author. While conventional teachers can resolve any open questions in class, (slicing) book authors have to foresee and solve all potential open issues in advance. This additional effort may make the preparation of additional material in sliced format unpopular, because the teacher author typically does not get the benefits of the primary author (e.g. the economic benefits of a wide-spread publications).

3. The underlying notion of learning by reading books or listening to lectures is incomplete. Modern didactics propose phases of active engagement with the content presented. By allowing learners to explore the content by asking questions and by presenting exercises and examples in a targetted manner, slicing books advance the state of the art of textbooks. Still, active engagement may require more interaction.

These limitations can be resolved if one regards slicing books as an active learning material rather than just an augmented book. An active learning material can be manipulated by all learning partners, the primary authors, the teachers and the students. By giving (some) authoring rights to the students much of the workload is transferred from the teacher to the students. The students benefit, because their authoring work can be framed as an active engagement with the content.

The remainder of this section will discuss this approach in the context of collaborative learning. Collaborative learning presents the learning content to the students in the form of a problem that they have to solve in collaboration. Typically they do so in subgroups using their analytical skills and creativity and by exploring and manipulating some prepared learning material. Collaborative learning is said to be greatly enhanced by computer-based tools supporting the learning process and by presenting more flexible computer-based learning material [WePf01].

How can slicing books be used as learning material for collaborative learning? The basic didactical approach builds on knowledge from moderation [Schw95, DeV99] and collaborative writing in a teaching context [BrBr97].

The lectures can take place in an electronic meeting room [Schw95] or the students bring their own notebooks along and use a wireless LAN provided by the University (Both exist in the new Metternich Campus of our University). There is a new division of labor: The teacher only supplies the basic slices of knowledge (plus some basic relationships). It is left to the students to further elaborate the slices, add new slices and to determine the relationships between the slices. The elaboration of slices gives the learning material the needed information at the needed level of detail. When the students determine the relationships between the slices they actively acquire the overview knowledge so often lacking in conventional lectures. While, in principle, each student could do these tasks on his or her own, additional value can be gained by regarding the working on a shared representation of the knowledge as collaborative task involving students and teacher. We call this approach "collaborative slicing" (ref. [DAFS01]).

Collaborative slicing can be seen as a special case of collaborative writing. Collaborative writing is difficult, because authors on the one hand need to be aware of what the others are doing, in order to produce a coherent result. On the other hand, they need times of undisturbed work to formulate their own thoughts. Typical collaborative writing efforts therefore distinguish between divergent (often asynchronous) modes, where each author works on his or her own section of work, convergent (and synchronous) modes, where all authors see the same content, and mixed modes where some authors work in a convergent mode and others in a divergent mode. Divergent phases further the deliberation of the content and assure that all needed information is included in the product; convergent phases assure that the product becomes a coherent agreed upon whole [DeVa99].

The collaborative learning scenario is based on divergent and convergent modes: The author prepares and present the content as unconnected slices in a divergent mode to a class of students. The students then split in subgroups of 2 or three persons and try to build the slicing relationships and thus explore the deeper structure of the content presented. In a third convergent phase the whole class is trying to arrive at a shared understanding of the slicing structure. There can be several variations of this approach: a) the student prepare the slicing relationships at home and only the subsequent phase three is done in classe. b) the author and the readers together discuss the slicing relationships in an oral discussion. Here content discussion and metadiscussion can be interwoven c) the author presents only parts of the information necessary. The readers then have to collect the other data and prepare the slicing hierarchy. This approach can be useful for case studies or experiments d) besides the shared representation of slicing book content and slicing relationships exist individual annotations and individual

slicing relationships. This approach would take individual preferences on knowledge representation into account. Current tools do not sufficiently support this collaboration yet. Requirements for a future collaborative tool include

- a) an easy to use interface for ad-hoc in-class slicing,
- b) shared access to a shared slicing representation (incl. concurrent access, locking, view linking, telepointing etc.)
- c) personal annotations to a slicing representation
- d) author-control of the collaborative process
- e) functionalities to test the slicing representation

We are currently experimenting with using a first primitive version of such a tool in the logics class.

7 Related Concepts

The concepts of Slicing Book Technology go considerably beyond that of hypertext. Especially, the generation of new customised documents from fragments of other documents – i.e. the concept "all information in one place" – is not found in hypertexts. Also note that the relations between slices can be of several types and can be used in combination with declarative descriptions of the desired documents. [DAFS01] will give a more thorough comparison of sliced books with hypertexts.

In contrast with efforts in the field of computer based training to compose personalised teaching materials from appropriate building blocks, Slicing Book Technology starts from pre-existing documents. This is the main reason why the technology is cost efficient. On the other hand, sliced books and the documents generated from them do not contain interactive assessment systems. Consequently, the Slicing Book Advisory System has to rely on the information provided by the user to build an internal model of the user's knowledge. Nevertheless, sliced books may augment the repository of materials of a computer based training system. The delivery tool developed within the TRIAL-SOLUTION project will have an open interface that allows external assessment systems to update the internal user model.

There are a number of projects in the academic world as well as on the commercial market that make use of disaggregated materials. Slicing Book Technology differs from these mainly in its extensive use of meta data.

The European project Ariadne (see [FWD97]) has set up a methodology to aggregate teaching materials from components. It has provided a system of meta data to describe educational features of teaching materials that has influenced several standardisation attempts.

Currently the European projects OR-World, 3DE and KOD investigate different aspects of the aggregation of personalized materials. In contrast with the approach presented here these projects work with materials that are specifically designed for this purpose.

A more general approach is taken by T. Berners-Lee [BHL01] and the W3C in considering the Internet as a semantic web consisting of semantic units with inference aided search procedures. In this sense a library of sliced books can be seen as a semantic web. Therefore we expect that our approach can help to gain early experience with the actual use of small semantic webs.

In the field of software reengineering a technique called Program Slicing [Tip95] has been developed. It disaggregates program source code and analyses the dependencies between software modules. This has the aim to identify and visualize the way in which the parts of a program influence each other. It can be applied to extract and reuse parts of a software system. A comparison of Program Slicing and Slicing Book Technology has been made in [La01].

It has been already mentioned in the introduction that electronic books might become interesting means to deliver personalized documents. Another form of delivery where this is already available now are print on demand services. With some of these services the user can online select parts of books – usually chapters or articles from journals – and gets a printed book delivered with exactly the selected content.

Unlike in Slicing Book Technology, the parts from which these books are composed are self contained and much larger than the slices of a book. Consequently, also personalized documents produced from sliced books can be much smaller. They are generated dynamically and instantly delivered over the web. The readers has the possibility to retrieve only what they actually needs, print this themselves and order a new document when their needs have changed. When they become familiar with this dynamic way to retrieve the requested information, this may even eliminate the need for a connection with a print on demand service.

8 References

- [BHL01] T. Berners-Lee, J. Hendler, O. Lassila: The Semantic Web, Scientific American, May 2001
- [BrBR97] R. Briggs, H. Brown: From the Sage-on-the-Stage to the Guide-on-the-Side: Re-engaging the Disengaged Learner with Collaborative Technology, unpublished manuscript 1997, <http://www.cmi.arizona.edu/personal/bbriggs/Downloads/index.htm>, download on June 11th 2001
- [CG77] Carr, B., Goldstein, I. Overlays: A theory of modelling for computer aided instruction (AI Memo 406). Cambridge, MA: Massachusetts Institute of Technology, AI Laboratory.
- [Da00] I. Dahn: Symbiose von Buch und Internet, Proc. Learntec 2000, Karlsruhe 2000
- [Da00a] I. Dahn: Automatic Textbook Construction and Web Delivery in the 21st Century, to appear in J. of Structural Learning Theory
- [Da01] I. Dahn: Slicing Book Technology - Providing Online Support for Textbooks, Proc. ICDE 2001, Duesseldorf 2001
- [DAFS01] I. Dahn, M. Armbruster, U. Furbach, G. Schwabe: Slicing Books – The Authors' Perspective. Unpublished manuscript prepared for: R. Bromme, E. Stahl: Writing hypertext and learning: Conceptual and empirical approaches, Pergamon press.
- [DeVa99] A. Dennis; J. Valacich.; Rethinking media richness: Towards a theory of media synchronicity. In: Sprague, R.: Proceedings of the 32th Hawaii International Conference of Systems Sciences (HICSS-32), IEEE Computer Society, Los Alamitos, California, et al. 1999 [CD-ROM of Full Papers, 10 pages].
- [Fu01] U. Furbach: Logic for Computer Scientists (Experimental Online Component)
- [FWD97] E. N. Forte, H. K. Wentland Forte, E. Duval: The Ariadne Project (part I) – Knowledge Pools for Computer Based & Telematics Supported Classical Open & Distance Education, European J. of Engineering Education, Vol. 22 No. 1, 1997
- [Ha97] M. Hazewinkel: Enriched thesauri and their uses in information retrieval and storage, In: C. Thanos (ed.), Proceedings of the First DELOS workshop, March 1996, ERCIM, 1997, 27-32
- [La01] C. Lange: Program Slicing and Slicing Book Technology – a Comparison, ARTI 8800, University of Georgia and University Koblenz-Landau, May 2001
- [Schw95] G. Schwabe: Objekte der Gruppenarbeit - ein Konzept für das Computer Aided Team, Gabler, Wiesbaden 1995.
- [Tip95] F. Tip: A Survey of Program Slicing Techniques, J. of Programming Languages, vol. 3 (1995), 121 – 189, 1995
- [WD00] H. Wolter, I. Dahn: Analysis Individuell, Springer-Verlag, Heidelberg 2000
- [WePf01] M. Wessner; H. Pfister: Kooperatives Lehren und Lernen. In: Schwabe, G.; Streitz, N.; Unland, R. (eds.): CSCW-Kompodium – Lehr- und Handbuch zur computerunterstützten Gruppenarbeit. Springer, Heidelberg et al. in print.